Smith Normal Form

Let R be a domain and $A \in R^{m\times n}$. We say that A is in *Smith Normal Form* (or SNF) if A is a matrix of the form

$$\begin{pmatrix} q_1 & 0 & \dots & & & & \\ 0 & q_2 & & & & & \\ \vdots & & \ddots & & & & \\ & & & q_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & \end{pmatrix}$$

with $q_1 \mid q_2 \mid q_3 \mid \cdots \mid q_r$ in R and $q_r \neq 0$, i.e., A has entries $q_1, \dots, q_r, 0, \dots, 0$ on the diagonal satisfying $q_1 \mid q_2 \mid q_3 \mid \cdots \mid q_r$ in R and $q_r \neq 0$ and all entries off the diagonal are 0.

Let R be a domain and $A \in \mathbf{M}_n(R)$. We say that $A = (a_{ij})$ is an *elementary matrix* of

i). *Type I:* if there exists $\lambda \in R$ and $l \neq k$ such that

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \\ \lambda & \text{if } (i,j) = (k,l) \\ 0 & \text{otherwise} \end{cases}$$

ii). *Type II:* if there exists $k \neq l$ such that

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \neq l \text{ or } i = j \neq k \\ 0 & \text{if } i = j = l \text{ or } i = j = k \\ 1 & \text{if } (k,l) = (i,j) \text{ or } (k,l) = (j,i) \\ 0 & \text{otherwise} \end{cases}$$

iii). *Type III:* if there exists a unit u in R and l such that

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \neq l \\ u & \text{if } i = j = l \\ 0 & \text{otherwise} \end{cases}$$

Remark. Let $A \in R^{m\times n}$. Multipliying A on the left (respectively right) by a suitable size

a.) Type I is adding a multiple of a row (respectively column) of A to another row (respectively column) of A.

b). Type II is interchanging two rows (respectively columns) of A.

c). Type III is multiplying a row (respectively column) of A by a unit.

Remark. All elementary matrices are invertible.

Theorem. *Let R be a euclidean ring, $A = (a_{ij}) \in R^{m\times n}$. Then A is equivalent to a matrix in Smith Normal Form (*SNF*). Moreover, there exist matrices $P \in \mathrm{GL}_m(R)$ and*

$Q \in \mathrm{GL}_n(R)$, *each a product of matrices of Type I, Type II, and Type III, such that* PAQ *is in* SNF.

Proof. The proof will, in fact, be an algorithm to find a SNF for A. Let δ be the euclidean function on R. If $A = 0$ there is nothing to do, so assume that $A \neq 0$.

Step 1. Choose $a = a_{ij} \neq 0$ such that $\delta(a)$ is minimal among all the $\delta(a_{lk})$, $a_{lk} \neq 0$. Put a in the $(1, 1)$ spot using matrices of Type II. In particular, we may assume that $a = a_{11}$. [If a is a unit in R, use a Type III matrix to make $a = 1$.]

Step 2. If $a|a_{ij}$ for all i and j, use Type I matrices to transform A into a matrix of the form

$$\begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

Note that a divides every entry of A_1 [Check]. If we are using the classical definition of euclidean ring every non-zero entry of A_1 has its δ value $\geq \delta(a)$. so δa is still minimal and we take A_1 and go back to Step 1.
[Note. If this occurs, by induction there exist matrices Q_1, P_1 such that $P_1 A_1 Q_1$ is in SNF. Let

$$P = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & P_1 & \\ 0 & & & \end{pmatrix} \text{ and } Q = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & Q_1 & \\ 0 & & & \end{pmatrix}$$

then

$$P \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} Q$$

is in SNF.] If we are using the modified definition of euclidean ring and a non-zero entry of A_1 has smaller δ value, $\delta(a)$ is not minimal so go back to Step 1. [As δa is a non-negative integer this cannot happen infinitely often.]

Step 3. Step 2 does not apply and there exists an entry $b = a_{ij}$ in either the first row or first column such that $a \not| b$:

Write $b = qa + r$ in R with $r \neq 0$ and $\delta(r) < \delta(a)$. Use Type I matrices to change A into a matrix with r in it. Since $\delta(a)$ is not longer minimal, go back to Step 1. [Since $\delta(a)$ is a non-negative integer and $\delta(r) < \delta(a)$, this must eventually stop.]

Step 4. Neither Step 2 nor Step 3 apply. Thus $a|a_{ij}$ whenever $i = 1$ or $j = 1$:

Use Type I matrices to convert A to

$$\begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

Now one of the following occurs.

(a) There exists a non-zero entry b in A_1 such that $\delta(b) < \delta(a)$. So $\delta(a)$ is no longer minimal. Go back to Step 1.

(b) $a|b$ for all entries b in A_1. This is impossible — You should have been in Step 2. [No matter, take A_1 and go to Step 1.]

(c) There exists an entry b in A_1 such that $a\not| b$:

Write $b = qa + r$ in R with $r \neq 0$ and $\delta(r) < \delta(a)$. Use Type II matrices to get b into the first colummn. (This does not change the $(1, 1)$ entry a.) Now use Type I matrices to change b to r. Since $\delta(a)$ is no longer minimal, go back to Step 1.

Clearly this algorithm yields a SNF of A. □

Remark. Note that we do not really need Type III matrices in the above. We only used them to transform those diagonal entries q_i that were units into one.

Remark. We did not really need Step 2 as it is incorporated into Step 4, but it is useful to use it as the euclidean rings that arise are usually the classical ones.

Remark. The SNF derived by the algorithm is, in fact, unique (up to associates). We will show this later.